U.S. Department
of Transportation

**National Highway
Traffic Safety
Administration**

People Saving People
http://www.nhtsa.dot.gov

DOT HS 808 466

October 1996

Executive Summary

Heavy Vehicle Driver Workload Assessment

Technical Report Documentation Page

1. Report No.	2. Government Accession No. D O T HS 808 467	3. Recipient's Catalog No.
4. Title and Subtitle Final Report--Program Executive Summary: Heavy Vehicle Driver Workload Assessment		5. Report Date September, 1995
		6. Performing Organization Code
7. Author(s) Louis Tijerina		8. Performing Organization Report No.
9. Performing Organization Name and Address Battelle Memorial Institute 505 King Avenue Columbus, OH 43201-2693		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTNH22-9 1 -C-O7003
12. Sponsoring Agency Name and Address Office of Crash Avoidance Research National Highway Traffic Safety Administration 400 7th Street, NRD-50 Washington, DC 20590		13. Type of Report and Period Covered Final Report July, 1991-September, 1995
		14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representative: Dr. Michael J. Goodman (NRD-50)

16. Abstract

This report summarizes a program of research to develop methods, data, and guidelines to conduct heavy vehicle driver-oriented workload assessments of new, high-technology, in-cab devices. Many such devices are being developed and implemented in heavy trucks and cars. Examples include navigation systems, text message display systems, and voice communications systems, to name a few. The objective of this research was the development of methods to assess the degree to which in-cab device use competes with the primary task of safely controlling the vehicle at all times. The following seven tasks, conducted throughout this program, are summarized: reviewing task analysis data and protocols literature; defining standard heavy vehicle configuration and tasks; collecting original task analysis data; reviewing workload measures and related research; developing a workload measurement protocol document; collecting baseline data of workload measures; and evaluating two high-technology systems using the develop protocols. Presented for each task is a summary of the objectives, approach, and key results highlights. From this research, tentative heavy vehicle workload assessment measures and methods were recommended and a protocol document was prepared.

17. Key Words ITS, heavy vehicles, human factors, driver workload, test and evaluation methods, guidelines	18. Distribution Statement

19. Security Classlf. (of this report)	20. Security Classif. (of this Page)	21. No. of Pages 44	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 PROGRAM OVERVIEW

In recent years, a wide variety of products have been proposed and developed for use in heavy trucks. These systems include the following:

- Satellite tracking, land navigation, and route guidance systems
- Text displays (e.g., pick-up address, package type)
- Vehicle subsystem monitoring and warning systems (e.g., tire pressure, oil pressure, brake failure, loading shifting)
- Computerized trip recorders (e.g., automatic record of speed, RPM, stops; driver entry of fuel purchase; state-line crossings)
- Sophisticated communication links (e.g., cellular phone systems)
- Proximity warning systems (e.g., infrared and TV systems)
- Portable personal computer technology (e.g., laptops, fax/modem), and
- Changes to existing control and display systems (e.g., head-up displays).

The heavy vehicle driver's primary task is to safely control the vehicle at all times. Many high technology devices introduce subsidiary tasks which may compete with the primary task of driving. This "competition" is what is meant by the term "driver workload" in this report. Some of these devices can probably be used concurrently with the primary driving task without interference, but others may not. It is reasonable to assume that the inventors and manufacturers of these systems intend for these systems to enhance commercial vehicle operations efficiency and effectiveness, to help the driver in doing the job at hand, and to be safe. However, without a driver-oriented assessment of a high technology device, the safety of the system remains largely unknown.

What is needed is a set of techniques with which to assess the safety implications of a device from the driver's perspective. In response to this need, the National Highway Traffic Safety Administration (NHTSA) initiated the program of research entitled, "Heavy Vehicle Driver Workload Assessment". In particular, the goal of this program was the development of a heavy vehicle driver workload assessment protocol. It is intended that the workload assessment protocol can serve as a basis for standard practice in the field of driver human factors test and evaluation. In industry, there exist Good Laboratory Practices (GLP), Good Manufacturing Practices (GMP), and IS0 9000 standards and certification. The field of driver-oriented test and evaluation of devices also benefits from similarly promulgated "Good Ergonomic

Evaluation Practices". The workload assessment protocol can serve as a draft for such a standard.

NHTSA contracted with Battelle, its subcontractor R&R Research, Inc., and several consultants to execute this program of research.. The objective of this research program was to develop methods, baseline data, and guidelines to evaluate the effects of high technology in-cab systems on a driver's ability to safety carry out the primary task of driving. The program of research was comprised of the following tasks:

- Task 1: Task analysis data and protocol review
- Task 2: Define standard heavy vehicle configuration and tasks
- Task 3: Task analysis data collection
- Task 4: Review of workload measurement and related research
- Task 5: Develop workload measurement protocols
- Task 6: Collect baseline workload data
- Task 7: Evaluate 2 high-technology systems.

A summary of each task is provided in subsequent sections of this executive summary. Each task also resulted in Task reports. A total of 9 volumes (including this executive summary) were finalized under the contract. These are listed at the end of this executive summary.

This research program included multiple studies which involved collecting data from professional heavy vehicle drivers. While a part task simulator was used for some aspects of the research, the emphasis of the research program was an empirical evaluation of heavy vehicle operation with professional drivers in an actual heavy vehicle on the road. As will be seen from a review of the individual tasks, dozens of drivers, hundreds of miles on the road, and a great deal of data collection were used to evolve the guidelines, methods, and results that are the products of this research.

2.0 TASK 1: TASK ANALYSIS DATA AND PROTOCOLS LITERATURE REVIEW

Standard human factors practice is to conduct a task analysis that describes human activities associated with the system under investigation. Thus, the objective of this task was to review available task analytic data and protocols pertinent to heavy vehicle operation and determine the availability and relevance of such data to heavy vehicle driver workload assessment. In addition, a preliminary consideration of the development of safety-relevant criteria was pursued as well as review of the relationship between risk-taking behavior and workload.

The task analysis data uncovered in Task 1 included American, Canadian, and European sources. The task analyses varied substantially in format and content. Some task descriptions were either too global or addressed driving conditions rather than tasks per se (e..g, "drive at night"). Other task descriptions were highly detailed and were prescriptive in nature. That is, they described the sequence of driver behaviors which should be executed rather than those that are actually executed. By and large, the task analyses were oriented to support training or certification. The task analysis data in the Task 1 interim report was used as the input for identification of standard driver tasks in Task 2 of this program of research.

A variety of protocol techniques were identified and reported on in the Task 1 interim report. Examples included activity analysis; interviews and commentary driving or protocol analysis; the critical incident technique; subjective workload ratings; visual allocation measures; on-the-road driver-vehicle performance monitoring; and safety criticality ratings and rankings. Many of these techniques would be incorporated into the execution of Task 3 of this program of research. Variants of these techniques would also find their way into the workload assessment protocol developed in Task 5.

The Task 1 report also introduced the thorny problem of establishing safety-relevant workload criteria. It was noted that no fully-developed methodologies or criteria were found with which to predict accident rates based on workload level. An actuarial approach, originally put forth by Perel (1976), was proposed. Perel's goal was to determine which in-vehicle systems were associated with higher accident rates using the detailed crash files in a North Carolina data base. Though subject to reporting inaccuracies, such data might be the basis upon which to predict accident rates based on "diversion of driver resources" or driver workload. It was suggested that visual attention or allocation measures, estimates of which could be derived from the literature or field tests, could be applied in a multiple regression framework to provide estimated or fitted crash incidence using the actual crash records as the "observed" values of crash incidence. This approach would later be elaborated on in Task 4 of this program of research.

Other approaches were also introduced to relate workload to highway safety. Visual allocation measures were introduced. To assess an in-cab device, mean single device glance duration, mean number of glances, total glance time away from the road scene, and mean single road glance duration during task execution were listed as specific visual allocation measures of interest. Coupled with the actuarial approach, visual allocation measures could be used to establish an empirical link between driver workload and safety. Lanekeeping measures such as lane exceedences were also introduced as potentially valuable measures that were safety-relevant. Lanekeeping measures would later play a prominent role in the measurement system for driver workload.

There is little doubt that some degree of risk is associated with virtually all activities engaged in by operators of motor vehicles. This suggests that there may be an important relationship between risk taking and workload. The Task 1 interim report examined the scant research literature on this connection and the following conclusions were reached:

- The concept of risk is established only at a theoretical level in terms of driver behavior. From a psychological standpoint, the existing work is weak and disjointed. Empirical work is both scarce and not tied well to theory.
- There are no established measures for assessing risk-taking in driving behavior.
- The relationship between workload and risk taking is not established in the literature reviewed. Thus, there was no justification to expect risk taking to raise or lower in the face of in-cab device-induced workload (or vice versa).

The difficulties associated with determining a link between risk-taking and workload was seen as very great. This topic would be addressed again in Task 4 of the heavy vehicle driver workload assessment project.

3.0 TASK 2: DEFINE STANDARD HEAVY VEHICLE CONFIGURATION AND TASKS

The objective of this task was to identify a standard heavy vehicle configuration and driver tasks that would serve as baseline conditions for measuring workload. The approach taken to identify the vehicle configuration was to make use of a subject matter expert in heavy vehicle operations and to review information on key states contained in the Department of Transportation's ***Truck*** Inventory and Survey. The approach taken to identify a set of standard driving tasks was to make use of the task analysis data provided from Task 1 of this program of research. The project subject matter expert consolidated the information into a single comprehensive set of driver behaviors and tasks. The subject matter expert then identified those tasks that were judged to be most relevant to the evaluation of in-cab device interaction.

The standard vehicle configuration considered most common was determined to have the following functional characteristics:

- Combination tractor and single trailer vehicle
- Conventional cab configuration with sleeper box optional
- Flat panel dashboard
- Diesel powered with air brakes
- Absence of high-technology devices

These functional characteristics were applied in the selection of a heavy vehicle for the project with which to collect workload-related measures on the road.

The identification of standard driving tasks began with a review of task listings and descriptions compiled in Task 1. These were consolidated into a set of conceptual categories and rewritten as needed in language that would be familiar to drivers. The heavy vehicle operation subject matter expert then identified key driving tasks based on many years of experience in all aspects of driving (from dispatching to vehicle operation to safety analysis to driver training). Table 3.1 presents the list of standard driving tasks when the vehicle is in motion. Items with asterisks represent those driving tasks thought to be most germane to driver workload assessment. To complement this table, the Task 2 report contains the following list of basic driver behaviors (sub-tasks) which lead to the completion of tasks given in Table 3.1:

- Looking at the road scene through the windshield

- Turning the head to view either west coast (side) mirrors
- Glancing down at gauges or controls (e.g., instrument panel)
- Turning the steering wheel
- Holding the steering wheel steady
- Moving the transmission gear selection lever
- Moving the accelerator pedal
- Moving the brake pedal
- Moving the clutch pedal
- Manipulating dashboard controls
- Adjusting the driver’s seat for comfort
- Adjusting windows for proper ventilation
- Adjusting air conditioning vents for comfort.

Many of these driver in-cab behaviors would later be used to develop candidate driver workload measures.

Table 3.1 Proposed Standard Driving Tasks (Source: Turanski and Tijerina, 1992).

Conceptual Categories	Associated Driving Tasks
Basic Driving Tasks	Start vehicle in motion Shift gears Reach desired speed in each gear Reach desired cruise speed * Control truck speed to allow for safe stopping distance * Brake under normal circumstances * Maintain safe following distance * Control direction via the steering wheel * Maintain lane position and spacing, straight road * Be aware of changes in the road scene [the primary visual task] Glance at gauges * Glance at mirrors Drive on a downgrade (steep gradient) Drive on an upgrade
Parking and Related Activities	Park tractor-trailer Back-up
Lane Changes and Passing/Overtaking	* Change lanes Pass on the left, cars (multi-lane, divided road) Pass on the left, other trucks (multi-lane, divided road) Pass on the left, cars (two-lane, undivided road) Pass on the left, other trucks (two-lane, undivided road) Pass construction zones * Merge Exit using an exit ramp
Turns and Curves	Make a left turn Make a right turn * Negotiate a curve and remain in your lane * Negotiate a curve and change lane in a multi-lane divided highway Turn your tractor-trailer around
Intersections and Crossings	Travel through intersections (You have right-of-way) Stop at intersections (They have right-of-way) Start truck in motion from a stop at an intersection Cross railway grade crossings * Negotiate l-lane and narrow 2-lane bridges * Negotiate narrow lane tunnels Stop at and start from narrow-lane toll plaza

Table 3.1 (Continued)

Conceptual Categories	Associated Driving Tasks
Nonstandard Driving	Recover from locked brakes due to extreme loss of air pressure Make a quick stop (Put a lot of pressure on brakes, but with no smoking tires, no danger of losing control) Make a hard braking stop (smoking tires, danger of losing control) Stop due to lighting problem (e.g., trailer lights go out) Stop due to engine problem (e.g., high engine coolant temperature, low oil pressure) Recover from tire failure, front tire(s) Recover from tire failure, other tire(s) Steer to avoid something on the road Recover from a tractor/trailer skid Respond to cargo or tire fire Execute off-road recovery (veer off the road to avoid collision, then immediately return to roadway

4.0 TASK 3: TASK ANALYSIS DATA COLLECTION

The objective of this task was to collect original task analytic data to support heavy vehicle driver workload assessment and protocol development. Data was collected from professional drivers to provide insights into the following issues:

- the meaning of the term "workload" to heavy vehicle drivers (N = 41 truck drivers interviewed);

- the demand placed on the driver by various driving condition factors (N = 55 truck drivers participated in study)

- the safety criticality and difficulty of selected standard diving tasks (N = 30 truck drivers interviewed);

- the perceptual, motor, and cognitive loads imposed by various tasks while on the road under various driving conditions (N = 9 truck drivers observed on the road during revenue runs); and

- the features of selected high-technology in-cab devices and loads that might be placed on heavy vehicle drivers.

Each of these component task analyses will be described in terms of approach and the key findings will be presented.

To determine the meaning of the term "workload " to heavy vehicle drivers, i.e., what factors are involved, a total of forty-one truck drivers were recruited from a local (Columbus, Ohio-area) truck stop. Each driver was interviewed one-on-one to discuss the factors that contributed to workload. The interviews consisted of open-ended questions and driver ratings and rankings of selected items. The interviews lasted from approximately 12 to 20 minutes. Key findings are as follows. When professional truck drivers think of "workload", they more or less consistently speak of time stress or stress caused by delays to their schedules. Thus, schedule delays and operational practices are most prominent in the minds of truck drivers in creating high "workload". City driving in traffic, heavy traffic, road construction zones, and inclement weather were most often mentioned as creating with high workload. The results from this investigation provided inputs into other activities in Task 3. Not surprisingly, the drivers most often reported that they compensated for increased driving workload by paying more attention to their driving task.

The results of the initial data collection indicated that, apart from the global impact of schedule delays and operating practices, primary driving task demand is determined largely by driving conditions outside the cab of the heavy vehicle. In a separate analysis, a psychological scaling approach was taken to determine the relative importance of five driving condition factors on the demand placed on the driver by the driving task. This approach allowed a unidimensional scale to be constructed with various combinations of these factors positioned along it. Based on data collected from fifty-five professional drivers, in decreasing order of their impact on demand, results were:

- traction,
- visibility,
- traffic (density),
- road type (divided, undivided roadway), and
- lighting (day, night driving conditions).

Traction and visibility were weighted consistently and considerably more important to the level of driving condition demand a driver faces. This suggested that the factors that contribute to the greatest driving task demand are those (traction, visibility) in which the driver has the least control over the vehicle. Safety considerations would later dictate that traffic, road type and lighting be incorporated into on-the-road workload testing; inclement weather, which gives rise to poor traction and reduced visibility, was judged too risky in which to conduct workload data collection.

To evaluate the safety criticality and difficulty of various driving maneuvers, thirty professional truck drivers were interviewed from a local truck stop. These drivers drove a variety of rigs and tractor types, and averaged approximately 15 years in driving experience. Each driver reviewed 5 tasks taken from the Task 2 listing of standard driving tasks and judged the safety criticality and difficulty associated with their execution under three selected driving conditions chosen to span a range of driving demand. A total of 30 tasks were reviewed by the participant drivers. Results were consistent with the driving condition demand assessment pursued with the psychological scaling approach, thus providing some validation of those scaling outcomes. The most difficult and safety critical tasks tended to be those that required the greatest visual/perceptual demand and precise or quick steering/braking control. This would later support the emphasis on visual allocation measures and steering/braking control inputs in the workload protocol developed in Task 5 and applied in Task 6 and Task 7 of this research program. It also indicated driving maneuvers that would not be incorporated into the on-the-road data collection because of safety considerations.

The first on-the-road data collection, with nine heavy vehicle drivers on revenue runs, was conducted to gain further insights into how professional drivers allocate their visual attention. Visual allocation and manual activity were captured by means of video tape recording techniques, Natural glance distribution by location was captured as well as the average glance duration and mean number of glances required to complete selected “commanded” tasks, i.e., commonly executed in-cab tasks that were requested of the driver by an on-board experimenter. Example results from this early on-the-road data collection are presented in Table 4.1 and Table 4.2. Results such as these were taken as early indications of the sensitivity and usefulness of visual allocation measures and would later play a prominent role in driver vehicle workload assessment and data collection. It also provided data that suggested that use of common in-cab tasks could serve as useful baseline measures against which to compare new high-technology systems.

The last task analysis effort completed in Task 3 of this project consisted of a preliminary analysis of generic high-technology in-cab devices. Table 4.3 contains the key features considered in evaluating the following generic classes of technology:

- Voice Communications Systems
- Vehicle Navigation and Route Guidance Systems
- Single/Integrated Displays (Status Displays)
- Text Communication Devices (Text Message Systems).

Based on this review, it appeared that a combination of any two of the following systems would provide a wide range of workload measurement challenges: route guidance system, voice communication system, and text display system. The latter two system types would later be the focus of the Task 7 device evaluation study.

Table 4.1 Task 3 Results: Heavy Vehicle Driver Glance Distribution by Location While Driving on the Road

Natural Glances by Location - All Subjects Combined		
	Mean Glance Duration (seconds)	**Percent of Total Time**
Left Mirror	1.33	5.6
Right Mirror	1.47	3.2
Instrument Panel	0.93	5.0
Header (Up)	0.83	0.8
Road-Ahead	3.85	76.4
Road-Left	1.22	1.2
Road-Right	0.98	2.0
Left Other	1.36	0.6
Right Other	1.28	5.3
		100.0

Table 4.2 Task 3 Illustrative Visual Glance Data from Heavy Vehicle Drivers for Common In-cab Tasks Executed While on the Road

Illustrative Visual Glance Data - All Subjects Combined			
In-Cab Task	**Average Glance Duration (seconds)**	**Mean Number of Glances**	**Average Time Off Road (seconds)**
Left Mirror-Detect	1.44	1.29	1.85
Left Mirror-Discrimination	1.77	1.58	2.79
Read Exact Speed	1.50	1.45	2.18
Manually Tune Radio	1.33	11.31	15.10
Change CB Channel	1.18	3.93	4.63
Wipers On/Off	1.00	1.13	1.13

Table 4.3 List of Device Dimensions Used for Preliminary Analysis of Generic In-cab Technologies

In-cab Device Dimension	Relevance to Protocol Development
When used	Indicates if device is used pre-trip, en route, only while parked, etc.
Relevant forward motion during driving tasks	See Task 2 Final Report Supplement for standard driving tasks. Will help identify reasonable times to present, observe, or otherwise assess in-cab tasks.
Driver-initiated or Device-initiated interaction	Important to establish locus of control. Driver may better manage workload if device has driver-initiated interactions.
Loads placed on driver	Visual, manual, cognitive loads are essential measures of the load imposed by an in-cab device.
Type of cognitive tasks required	E.g., check reading, text display, keying input, etc. would be useful information for assessing possibility of interference with primary driving task.
Time required	Important to note if an interaction is measured in fractions of a second, seconds, minutes, or portions of an hour. This will provide guidance on the necessary and sufficient timing properties of candidate workload measures.
Error modes	What can go wrong while using a device and likely driver reactions. May provide indication of driver involvement with in-cab device transactions.
Task steps	Indicated, perhaps, by mode of operation. May eventually be useful for explaining the locus of a workload effect.

Table 4.3 (Continued)

In-cab Device Dimension	Relevance to Protocol Development
Perseverance effects	Device allows driver to readily break with task and come back to it versus device which prompts driver to persevere (e.g., because needed data are rolling off the screen, because system timeouts reset a control, etc.). This feature of devices is thought to have high relevance for highway safety.
Importance of interaction and driver discretionary device use	Perceived criticality (e.g., must do, optional); an indication of the urgency with which the driver will want to respond to the device.
Positioning	Where device is mounted; on dash, overhead, on seat, other. May affect driver posture and lead to loss of visual awareness of the driving situation.
Likelihood of use with other systems	What other systems might be used with the device in question. E.g., a multifunction/integrated digital communications system may not be used with a trip recorder (because it already has that function built into it) but could be used with, say, a cellular phone.

5.0 TASK 4: REVIEW OF WORKLOAD AND RELATED RESEARCH

Task 4 was primarily a literature review that focussed on workload measures and related research. It included an attempt to develop a theoretical basis for relating driver workload to highway safety. The team conducted a selective review of the literature on driver performance evaluation, workload evaluation in a driving context, and risk taking and risk adaptation. An actuarial approach to establish the safety relevance of workload measures was presented, along with a driver resource allocation model of in-cab device workload. Finally, a driver resources-based taxonomy of in-cab tasks was presented along with candidate workload measures and their potential sensitivity to tasks from the taxonomy.

The literature review supported the development of a view of the driver as an essentially rational individual who will act consistent with his or her current understanding of the driving situation and motivations. So it is expected that in-cab technologies will be used safely most of the time if the technologies are quick and easy to use. Otherwise, these technologies may not be used at all, unless operational practices make their use mandatory. Safety is compromised when the driver believes it is safe to work with a device while driving and, in fact, a safety hazard lurks on the roadway. It was assumed that the driver may be able to do several things at once but the driver can consciously focus attention on only one task at a time; generally, this focus is visual in nature.

Of several means to establish the safety-relevance of workload, the most promising was termed an "actuarial approach". As depicted in Figure 5.1, as indicated, a crash data base would be selected. In Step I, the selected data base would contain records of sufficient detail to determine the probable cause of the crash. The data base would also allow for searches by key words and phrases related to in-cab distractions or competition with the driving task (e.g., "attending to radio", "reading a map","checking instrument panel", etc.). A list of search terms would be applied in a data base search in Step II. The returns would subsequently be screened for relevance to driver workload (Step III). On a parallel track, two other kinds of safety-critical data would be sought out. One type of data would ideally provide estimates of the frequency of in-cab device use as an index of exposure to the in-cab distraction hazard (Step IV). The other type of data (Step V) would provide estimates of visual workload associated with the in-cab devices or distraction sources indicated in the data base. Given this type of information, the final step would involve building regression models to relate crash incidence to measures of driver workload and device frequency-of-use (Step VI). This work would later be carried out in the research program, the results of which are reported an appendix of the workload protocol drafted in Task 5.

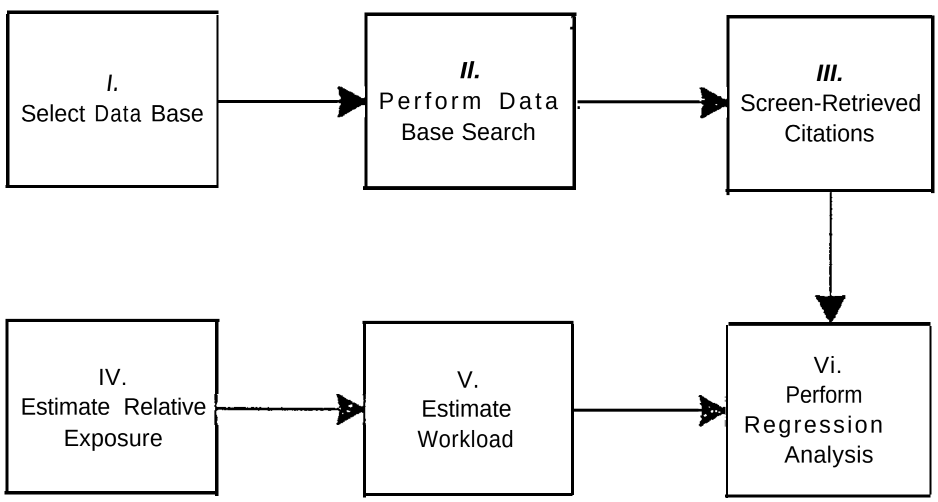

Figure 5.1 Block-diagram of the Steps Involved in Relating Crash Data to Workload

A simple model was presented that represented driver resources that must be shared between in-cab device use and the driving task (see Figure 5.2). The research literature to support this model was reviewed and it was determined that visual, manual, and cognitive demands were of primary importance to heavy vehicle driver workload assessment. Furthermore, the following driver resources-based taxonomy of in-cab tasks was presented:

- Manual Only — Tasks performed without visual reference, e.g., pressing "set" or "resume" on a cruise control;

- Manual Primarily — Vision used to find a control, then task performed without further visual reference, e.g., adjusting radio volume;

- Visual Only — Tasks are completely visual in nature and have no manual component, e.g., reading speedometer, clock, or air pressure gauge;

- Visual Primarily — Tasks that are predominantly visual but have some manual activity associated with them, e.g., determining radio station frequency when the radio display is initially set to display time;

- Visual-Manual — Tasks with interactive visual and manual demands, e. g.., manually tuning a radio, operating a cellular phone, zooming in and out of a map display; and

- Cognitive — Tasks with little or no visual demands or manual demands, e.g., dialogue on a cellular phone.

Task 5 also included a listing of driver-vehicle performance measures potentially sensitive to in-cab device workload. The likely sensitivity of these candidate workload measures were related to the task types in the taxonomy. The task taxonomy would later be used, along with other information provided by preceding tasks, to select requested tasks in Task 6 baseline data collection and device tasks for Task 7. The listing of candidate workload measures would be elaborated upon in Task 5 when drafiing the workload assessment protocol.

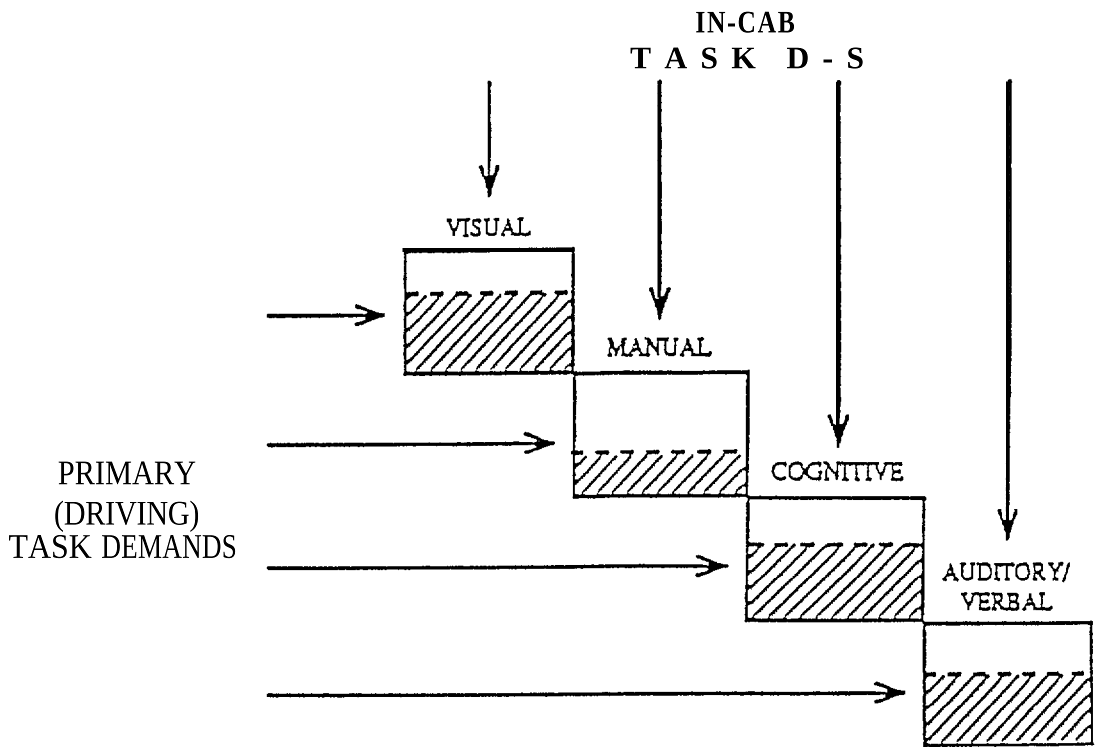

Figure 5.2 Depiction of the Competition for Driver Resources

6.0 TASK 5: DEVELOP WORKLOAD MEASUREMENT PROTOCOLS

Task 5 had as its principal objective the drafting of a workload measurement protocol document. At this time, a decision was made by NHTSA to direct the effort toward the development of protocols that emphasized rigorous, empirical data collection under realistic driving conditions. To this end, the protocol emerged to support workload assessment using instrumented vehicles and procedures that could be carried out on the road.

The primary objective of a workload assessment of an in-cab device or system is to empirically assess the potential of that device to distract the driver from the driving task to the extent that safety may be compromised. Given this primary objective, the workload protocol document describes a process by which such an assessment may be carried out. It is intended to be applicable to a wide variety of in-cab or in-vehicle devices. In addition, it is intended to support a wide range of individuals who are charged with the responsibility of assessing the distraction potential of new high-technology for use in heavy vehicles.

This wide scope necessitates a general document that provides guidance on the conduct of workload assessments. This document presents a series of stages which, if carried out, will promote a more thorough device evaluation. It does not, in general, provide a single, fixed evaluation procedure because variation in categories of technologies and their uses by drivers does not allow it.

This document is targeted to several potential users:

- The protocol document is intended to be of use for new or novice evaluation team personnel and for test engineers who may have little or no experience with driver-oriented data collection and assessment.

- The DOT/NHTSA may use it as a guidance document to manage contractors retained to carry out safety evaluations, especially operational tests. The steps/stages discussed in the document may serve as a set of milestones for a formal evaluation and ensure that all relevant factors have been addressed.

- The document may be of use to researchers in the field of driver workload. Experience has shown that there are special aspects of driver-centered device evaluation that are different from both psychological measurement or engineering assessments. For this reason, there is value in having a guide to the development and conduct of a workload evaluation.

While the protocol document is intended to be practical, it describes an idealized process of evaluation, i.e., it is prescriptive or tells what ought to be done. On the other hand, the rigors of realistic evaluation on a specific device or system in a specific circumstance may demand procedural changes that vary somewhat from what is described in the protocol. Each evaluation has its own unique qualities that may or may not be adequately expressed in general principles.

The protocol developed is intended to serve as a basis for standard practice in the field of driver workload test and evaluation. In industry there exist Good Laboratory Practices (GLP), Good Manufacturing Practices (GMP), and IS0 9000 standards and certification, to name but a few. The field of driver-oriented test and evaluation also will benefit from similarly promulgated "Good Ergonomic Evaluation Practices". The workload assessment protocol can serve as a draft for such a standard.

In an effort to provide a comprehensive document, the workload assessment protocol consists of several parts. It begins with a simple model of driving from which several categories of candidate workload measures are derived:

- visual allocation measures;
- driver in-cab behavior and control input measures (steering, accelerator, brake inputs);
- driver-vehicle performance measures (speed, headway, and lanekeeping measures); and
- subjective assessments.

Additional potential measurement categories (in-cab device performance measures, traffic conflict measures, and more abstract human factors measures) are introduced. Their exclusion from the workload assessment protocol is then explained.

Additional sections address the validity of workload measures, their scientific bases for the safety relevance of workload measures, and the prospects for successfully predicting the safety impacts of new, high-technology systems. It is in the protocol, for example, that the results of project activities to carry out the actuarial approach are presented. For example, Figure 6.1 presents a portion of the results obtained from the crash database search. It indicates the number of crashes distributed by sources of attentional distraction, broken down further into various interior and dash/console/steering column sources. This distribution of crashes became the "observed" data to estimated exposure and (visual) attentional demand would be applied to develop a prediction equation. Frequency-of-use data for various in-cab devices were gathered from the literature and through application of engineering judgment. The attentional demand estimates also came from the literature on visual allocation (mean glance durations and mean number of glances to complete in-cab tasks) as well as engineering judgement when necessary. These data were collated into a composite predictor variable termed "Exposure," and this was used in a regression framework to estimate crash incidence; the incidence of crashes gathered form the database search was used as the observed data. An example of the regression results is provided in Figure 6.2. The regression models generally provided an excellent tit between estimated and actual crash incidence given an exposure value that was the product of mean single-glance time to a device or location, mean number of glances to that device or location, and device or location frequency of use.

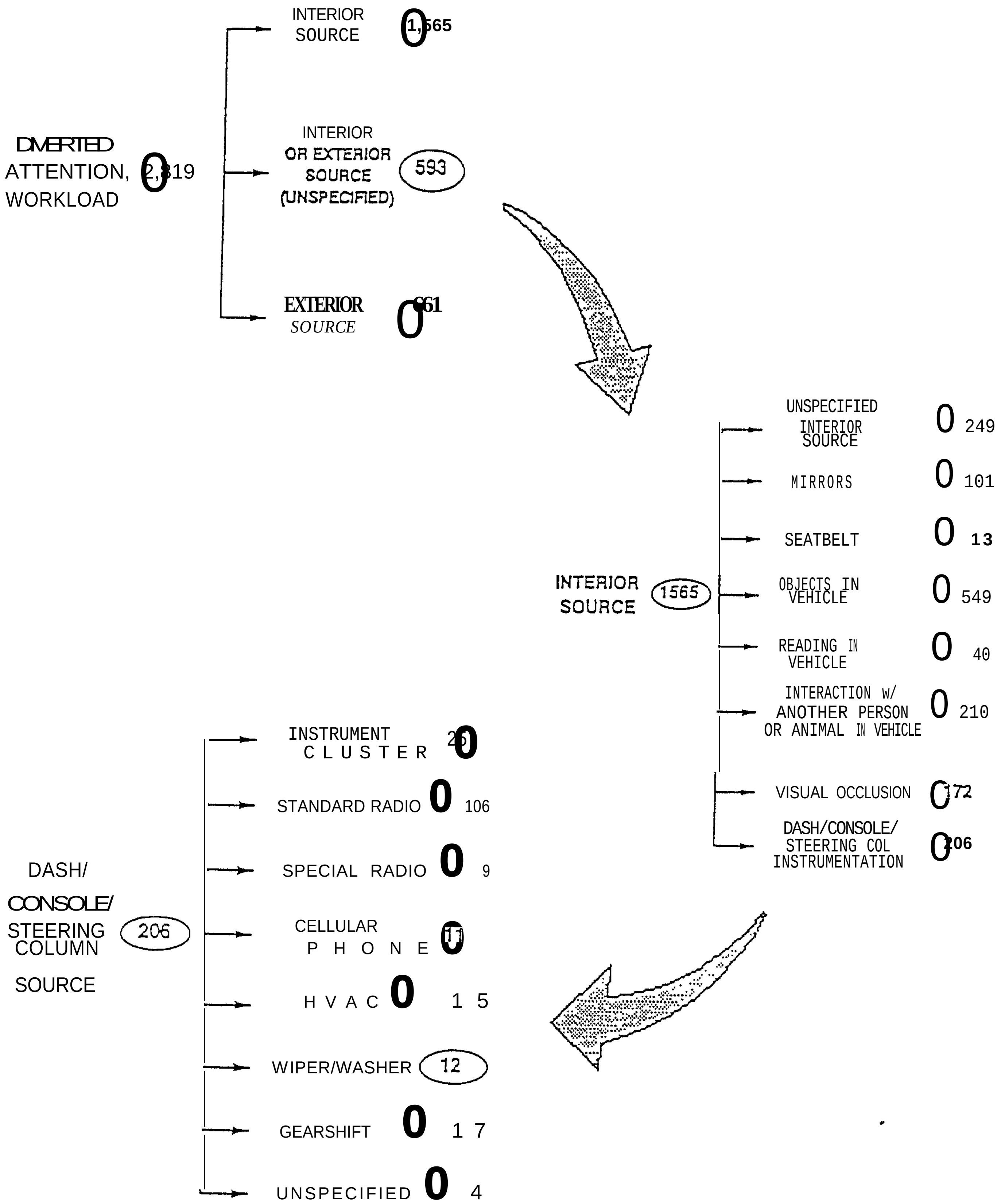

Figure 6.1 Number of Crashes Distributed by Sources of Attentional Distraction, Broken Down into Interior Source and Dash/Console/Steering Column Instrumentation Group

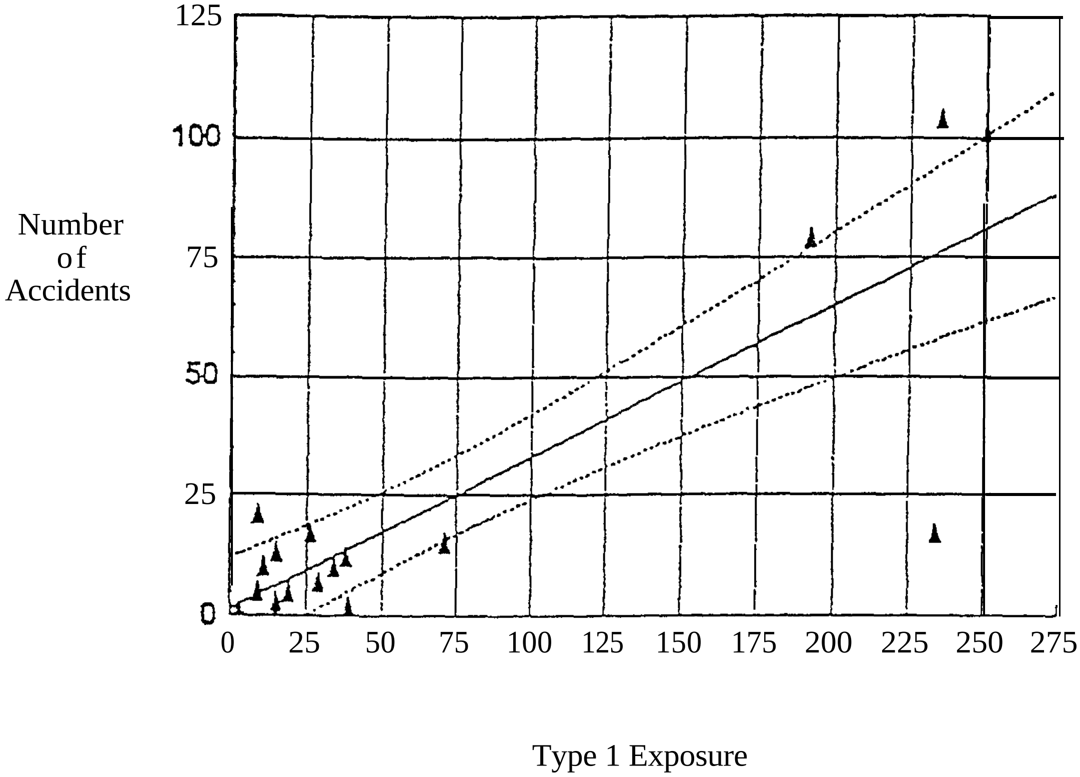

Figure 6.2 Plot of Crash Incidence vs. Type I Exposure, Defined as the Product of Average Single Glance Time to a Device, Mean Number of Glances to a Device, and Frequency of Use

Figure 6.3 presents the overall workload assessment protocol process. Its complexity is apparent. However, every attempt was made to be comprehensive. Furthermore, it is expected that the process or protocol developed for a specific assessment will tailor the steps accordingly. The document includes discussions of each phase and step of the process and provides guidance on the execution of a rigorous, empirical workload assessment based on an instrumented vehicle and on-the-road data collection.

The protocol document concludes with several appendices that bring together operational definitions of candidate workload measures, literature reviews on their use in past research, and discussion of the instrumentation needed to collect and process such measures. Figure 6.4 presents an excerpt from one appendix that defines, both graphically and in narrative, a steering hold and its expected workload interpretation.

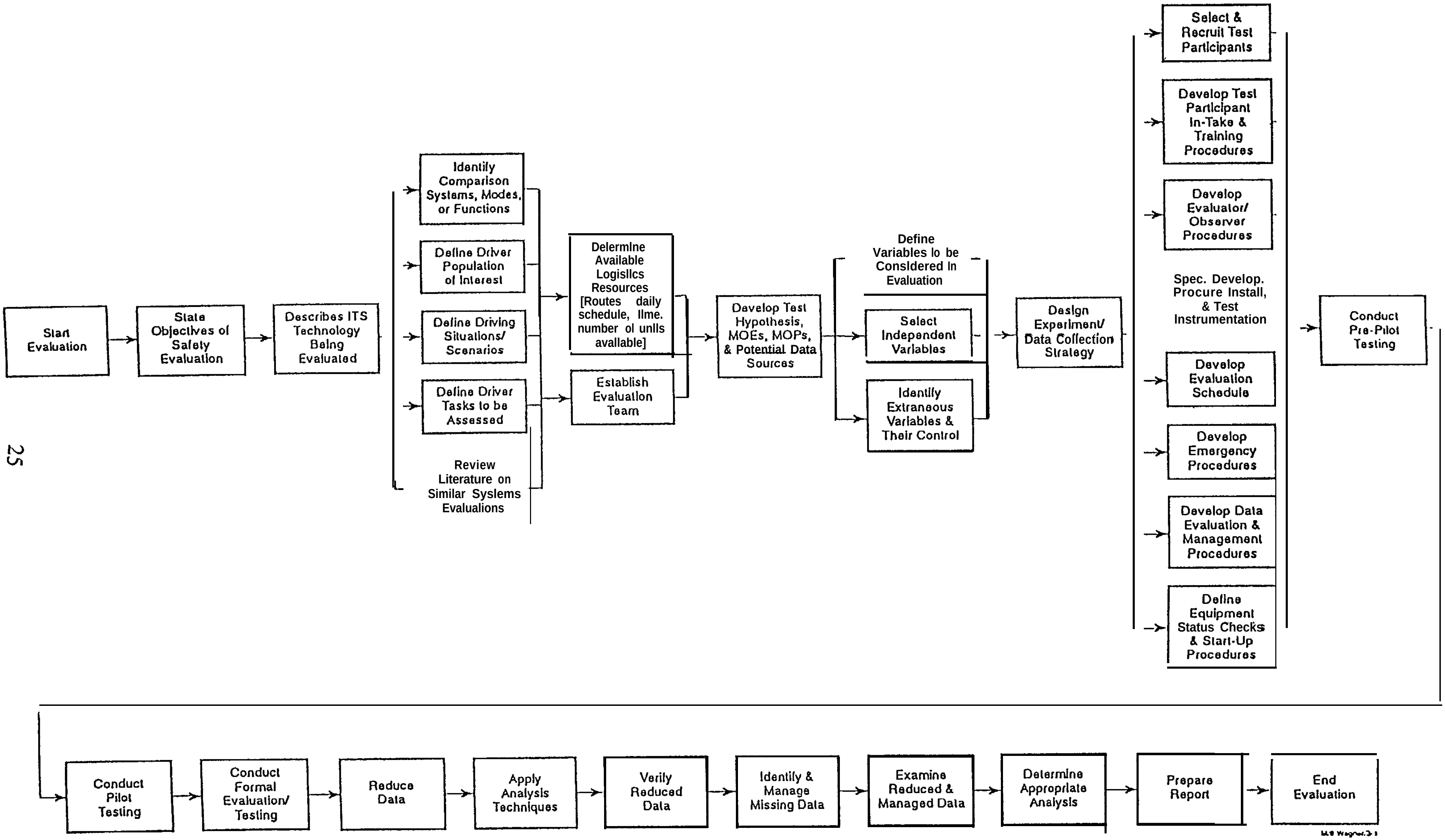

Figure 6.3 Flow Diagram of the Driver Workload Assessment Protocol

Number of Steering Holds = Total number of holds within a sample time interval.

Workload interpretation: If in-vehicle device demand is high, the driver will have to direct his or her attention to the device, numerous times. During such periods, the driver may hold the wheel relatively still, then make a corrective input after taking a glance to the roadway. Thus, the number of steering holds may increase as task demand increases.

Note that number of steering holds and steering hold duration may trade off within a fixed sample interval. That is, very long hold durations may be indicative of high workload demand, yet may be associated with fewer rather than more steering holds. Thus, it is important to consider the two measures together, especially if the sample interval is fixed rather than allowed to reflect task completion time.

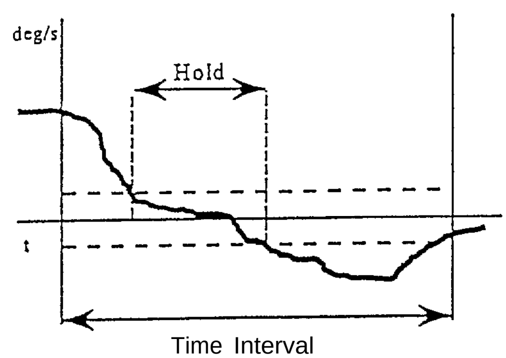

Figure 6.4 Excerpt from the Driver Workload Assessment Protocol: Definition of Steering Holds and Workload Interpretation of Number of Steering Holds

7.0 TASK 6: BASELINE DATA COLLECTION

This was the sixth in a series of tasks involving the assessment of driver workload in heavy vehicle operation associated with in-cab devices or systems. This phase of the work had as its chief objective the development of a baseline of driver visual allocation, in-cab behaviors, and driver-vehicle performance under different driving conditions, driving tasks, and in-cab tasks while on the road. The measurement system of candidate workload measures included in the workload assessment protocol served as a guide for the measures taken in Task 6.

Each of thirty licensed truck drivers drove an instrumented heavy truck for approximately 4 hours over a 459 km fixed route in which road type (urban freeway, rural freeway, Z-lane rural road), and ambient light levels (night driving, daylight driving) were varied. In addition, both open road driving and car following scenarios were observed. Finally, requested tasks, i.e., commonly executed in-cab activities, were included to determine their efficacy as baseline conditions for future workload evaluations and research.

Results indicated that visual allocation measures were especially sensitive and robust to driving conditions and requested tasks. Lane keeping measures and speed measures were also sensitive to variations in driving conditions and requested tasks. Of the driving condition factors, road type had the most pronounced effect. Light effects were minimal, as were driving scenario effects when comparing open road driving to car following scenarios. This is thought to be at least in part an artifact of the experimental procedure because traffic density was sparse and no requested tasks were executed if there was a lead vehicle within approximately 61 meters (200 feet) of the heavy vehicle. For this reason headway measures were largely absent of effects.

Table 7.1 presents a sample of the visual allocation data collected in Task 6. In general, Task 6 led to the following conclusions:

- The Task 6 report contains a comprehensive look at professional heavy vehicle drivers driving a standard tractor-trailer rig with a single trailer containing approximately 34,473 kg of dry freight. These data characterize driver workload under nominally "normal" driving conditions;

- The Task 6 report indicates that while driving condition factors such as road type and lighting can influence candidate workload measures, they are, by and large, additive in their effects with requested or in-cab tasks. This is beneficial in that it simplifies interpretation of results. However, driving conditions should nonetheless be taken into consideration in driver workload evaluation. Note that safety considerations may limit the range of driving conditions that can be examined in a workload assessment on the road. For example, even though Task 3 results indicated that poor traction and visibility contributed most to driving demand, safety considerations precluded data collection in inclement weather.

- Task 6 results confirmed that, as baselines, open road driving and manual radio tuning have merit in terms of characterizing commonly incurred "minimal" and "maximal" workload. Thus, these may be used as lower and upper anchor points with which to gauge the workload imposed by new devices in a relative workload comparison.

Table 7.1 Visual Allocation of Heavy Vehicle Drivers on the Road, by Location

GLANCE DURATIONS			
Glance Location	N Obser-vations	Mean Glance Duration(s)	Glance Duration(s) SD
Road scene-natural	689	2.83	2.43
Off road-natural	689	0.93	0.21
Left mirror			
Natural	618	0.97	0.29
"Requested" (left mirror-detect)	217	1.21	0.55
Right mirror			
Natural	403	0.96	0.33
"Requested" (right mirror- detect)	230	1.37	0.59
Inst. Panel			
General	610	0.84	0.23
Digital Clock*	231	1.20	0.43
Air Pressure Gauge*	238	1.57	0.71
Radio*			
Tune Radio	228	1.22	0.41
Adjust Volume	239	0.77	0.40
Tune CB*	233	0.96	0.34

NUMBER OF GLANCES TO COMPLETE "REQUESTED" TASKS			
Requested Task	N Obser-vations	Mean Number of Glances	Number of Glances SD
Read Digital Clock	231	1.03	0.17
Read Air Pressure	238	1.16	0.46
Adj. Radio Volume	239	1.10	0.54
Tune Radio	228	5.62	3.15
Left Mirror-Detect	217	1.05	0.28
Right Mirror-Detect	230	1.05	0.22
Tune CB	233	3.23	1.33

*Requested.

8.0 TASK 7: EVALUATE TWO HIGH-TECHNOLOGY SYSTEMS

Task 7 of the program of research involved evaluation of two high-technology systems using the developed protocols. The objective of this task was to apply the workload assessment protocol and measurement system to the selected technologies and determine characteristics and conditions of implementing those technologies that can have undesirable safety consequences. The results of the device use would be compared to baseline workload measures taken to provide indications of the driving conditions and device characteristics under which it would be safe to use or operate the technologies.

A prototype text message display system and cellular phone system were chosen for evaluation. These choices were based on work from Task 3 that indicated such systems would provide a range of workload assessment challenges. Furthermore, the in-cab task taxonomy developed in Task 4 indicated that these two technologies would involve a range of tasks with which to determine the sensitivity of candidate workload measures. These two technologies also serve as surrogates for a broader range of technologies with similar attributes or driver demand characteristics. Thus, results from this study can be used to gauge the workload impact of systems with similar attributes. Note that text message display use would be a Visual Only task; manual dialing of the cellular phone would be a Visual-Manual task; dialogue carried out over the cellular phone would largely constitute a Cognitive Only task.

The study assessed the driver workload imposed by a text messaging system and cellular phone on heavy vehicle drivers under various driving conditions. Sixteen (16) professional commercial vehicle operation (CVO) licensed drivers drove an instrumented heavy truck over a 4-hour period on public roads under various conditions of ambient lighting (day or night), traffic density (light or heavy), and road type (divided or undivided). Within driving condition combinations, various levels of text message reading, cellular phone dialing, radio tuning, and cellular phone dialogue were completed by the driver. Continuous measures were taken of visual allocation, steering and accelerator activity, speed maintenance and lanekeeping performance.

Results were presented in the Task 7 final report for the following three main components of the study:

- Reading text messages of various lengths (l-, 2-, or 4-lines) and with various content from a prototype CRT text message display;
- Manually dialing a cellular phone in any of three dialing configurations (auto-dial, 7-digit dial, or IO-digit dial), along with manually tuning a radio as a baseline; and
- Engaging in question-and-answer dialogue with either biographic questions or arithmetic questions on a cellular phone, along with open road driving as a baseline.

The text message display analysis indicated that 2-line and 4-line messages like those used in this study can have substantial effects on visual allocation (increased time looking away from the road scene, shortened glances to the road scene during message reading) and lanekeeping performance (e.g., greater incidence of unplanned lane exceedences). Thus, a recommendation is offered that text messaging systems be kept to a one-line display of perhaps 55 characters, with the display parameters of character size, luminance, polarity, and contrast set to levels established by ergonomics research. It was also mentioned that, though not pursued in this study, there is evidence that increased workload and comprehension demands may be associated with attempts to put too much information by displayed abbreviations into even the one-line display format. This is an area in need of further research. However, past research strongly suggests that abbreviations for data display (as opposed to data entry) will be fraught with problems. Such problems may be aggravated by the demands of concurrent driving.

For this study, the analysis of manual activity focussed on various levels of manual dialing of a cellular phone. Manual radio tuning was included as a baseline condition based on previous research. Results indicated substantial differences in visual demand associated with 7-digit and lo-digit dialing. There were no substantial effects of manual task on speed maintenance or lanekeeping performance. However, lane exceedences were observed on 27 percent of all trials. Given that the drivers were probably very vigilant since they knew they are participating in a study, this is disturbing. As a point of reference, only about 14 percent of the trials involved lane exceedences when drivers read a l-line "What is the current time?" CRT text message. Furthermore, the manual radio tuning task was chosen as a baseline because it represents a societally accepted, though high-demand task. However, manual radio tuning is itself associated with crashes, as demonstrated in the actuarial analysis in Task 5. Thus, even with professional truck drivers, it is advisable to streamline the manual dialing aspects of cellular phone dialing, either with auto-dial features that minimize the number of keystrokes required or perhaps a voice-dial feature that does not require manual input or eyes off the road at all.

Heavy vehicle drivers tend to spend a great deal of time engaging in dialogues on CB radios. This has the benefit of keeping the driver alert while driving. However, the present study showed that even relatively simple question-and-answer dialogues can have subtle effects on safety-relevant driving behaviors. In particular, the results indicated that visual scanning, as measured by mirror sampling, was cut by almost 50 percent on average when the driver was engaged in dialogue as compared to engaged in open road driving without dialogue. This is a potential cause for concern in that it represents an increase in crash hazard exposure. From this, one may hypothesize that such in-vehicle device workload does not degrade the highly over learned skills of driver-vehicle speed and lanekeeping performance among professional truck drivers. Instead, in-cab device workload decreases driver monitoring for hazards on the roadway.

It is interesting to consider also subjective assessments made by the drivers during the data collection debrief When subjects were debriefed after the run, they reported some difficulty in phone use but over half reported no difficulty in executing the tasks. When asked to rate workload imposed by the four road sections, i.e., divided (I-270) roadway- day, divided roadway-night, undivided roadway (SR-161)- day and undivided roadway - night, the subjects rated

SR-161 to have a higher workload than I-270 and night operation with a higher workload than daytime operation. Thus, SR-161 night workload averaged 39.7 on an overall workload scale of 100 vs. 12.3 for I-270 day. The absolute value of these scores are less important than the relative scores. When asked to rate the in-cab tasks in terms of workload the drivers gave average workload ratings ranging from 19.5 for the 3-digit dialing task to 27.3 for the IO-digit dialing task with CRT reading and cognitive load falling between these two values.

When asked to rate workload (again on a 100 point scale) for the highest workload road condition (SR-161 night) and the highest task workload (lo-digit dialing) the subjects gave a mean response of 56.9 suggesting an additive effect. Most of the drivers believed the in-cab tasks were realistic although many expressed some concern about phone usage. This latter response is understandable since few had cellular phone experience and none had used them in regular truck driving. Thus, drivers in this study were apparently aware of the potential hazards associated with the technologies to which they were exposed.

This study was intended to demonstrate the sensitivity of candidate workload measures to assess workload variation with two in-cab devices. In terms of establishing the sensitivity of the workload measurement system to variations in heavy vehicle driver workload, this study met this goal. However, the original intent of the task was to establish the conditions under which it would be safe to use or operate the technologies tested. Here, the study is equivocal. This is because the state of the art in driver workload assessment is such that relative assessments are feasible but absolute assessments that predict crash occurrence are not feasible. For a detailed discussion of the difficulties associated with predictive safety impact assessments, see the workload assessment protocol document.

9.0 DISCUSSION

This program of research established an association between workload and highway safety through a variety of means. However,. it was not possible to develop and validate a quantitative model to predict crash incidence as a function of workload measures. This safety prediction is difficult to achieve for several reasons which may never be fully resolved. Examples include the chaotic nature of many crashes, the need for better or more precise data on crash circumstances (e.g., crash file narratives that describe distractions associated with crashes), and even the properties of statistical models themselves. Because of these difficulties, workload assessment is best considered as a relative assessment of a device or task in comparison to other tasks or baselines. Establishment of workload measurement red-line values is infeasible at this point in time.

A number of important issues which bear upon workload measurement are listed below. Future research is needed to address these issues.

Necessity for Realistic Tasks vs. Laboratory Tasks. At least some of the results reported in the driving literature make use of highly artificial stimulus materials which are intended to maximize the size of an experimental effect. For example, cellular phone research has often made use of mental arithmetic and grammatical reasoning problems as test messages. Perhaps not surprisingly, these types of stimuli show an effect of "workload" on driver-vehicle performance. However, it is important to extend these results into the real world with message length, content, pace, and level of driver interaction which more closely resemble that which is common to relevant in-cab devices. To use the voice communications example again, it might be important to develop an approach where typical messages between dispatchers and drivers are recorded and are used to develop the stimulus materials which will be integrated into workload assessment protocol scenarios.

The Use of Baseline Measurements with Conventional Tasks. Requested tasks were used to collect baseline data on the task demands which are a part of driving in a conventional cab without new high technology in-cab devices. It seems appropriate to consider extending this notion to include, when feasible, other baseline "control" conditions as part of the protocol development effort. This could involve non-automated alternatives to the functions provided by an in-cab device, e.g. :

- use of a paper map instead of an electronic map;
- use of a pad and pencil to make notes instead of a video text display;
- use of the conventional radio instead of a special communications system for weather updates.

Collecting data on driver-vehicle performance both with and without high technology in-cab devices should provide useful assessment information, e.g., that the in-cab device takes no longer than a conventional means to accomplish the same function. This type of baseline might complement more general baseline measures on the use of visual, manual, or cognitive resources on a variety of tasks not directly related to in-cab device functionality.

There is a potential problem associated with using conventional in-cab tasks as a standard of acceptable or "safe" device loads. For example, use of a car radio or cassette player is potentially dangerous, and the actuarial analysis conducted for Task 5 was successful in correlating in-cab visual loads to crashes. This actuarial approach might support using conventional in-cab driver loads as benchmarks for safety when evaluating a new in-cab device. Alternatively, the most one might conclude from comparison of new in-cab device use with baseline data as described here is that the in-cab device is or is not imposing any driver demands beyond those which are already present in the fleet.

Criticality of Primary Task Performance Covariation with In-cab Device Use. In earlier tasks, emphasis was placed on the notion that in-cab devices can be judged to negatively impact on safety if primary task performance is degraded. Primary performance might not be sensitive to changes in workload yet sensitive measures are traditionally sought for workload assessment. In traditional laboratory approaches to workload, sensitivity of measures to workload manipulations led to the development of alternatives to primary task performance (e.g., unembedded secondary tasks). While sensitive measures are highly desirable in a practical heavy vehicle driver workload assessment protocol, if primary driving performance is unaffected under a reasonably broad range of driving conditions, then it would be very useful to be able to conclude that the device assessed had no deleterious effects on highway safety. On the other hand, Task 6 and Task 7 of this program of research showed that visual attention to the driving situation (as measured by average road scene glance durations and mirror sampling) can be reduced by in-cab task execution without necessarily disrupting lanekeeping or speed control. Thus, safety relevant driver-vehicle performance measures must include both primary driver vehicle performance measurements as well as visual allocation measures that reflect driving situation awareness. If either category of measurements indicates in-cab device competition, this should be taken as an impetus to make the device less intrusive.

Problems with Best-case/Worst-case Analysis. It is possible to develop a test scenario which any device will look bad in (the worst case) or any device will look good in (the best case). Between these extremes is a range of driving conditions which will vary by driver, season, route taken, etc. There does not appear to be a clear means to determine what might be a "fair" set of scenarios to test a device in other than representative ones. However, emphasis should be placed on determining under what conditions a system might intrude on the driver and compromise safety. Subsequently, the likelihood of these conditions occurring in the real world might be assessed and some estimate made of the criticality of workload measurements taken under such conditions. Further, one must also consider actual and perceived demand characteristics of a device. Devices that require immediate attention such as warnings may need to be treated differently in a protocol than devices which allow the driver to decide when it is safe to respond to device activation.

The Effects of Individual Differences It is a truism that workload is an interaction between a task (or set of tasks) and a driver, In interacting with an in-cab device, there will most likely be considerable individual differences among drivers and even within the same driver over a time course because of learning, fatigue, etc. Common statistical tests assess mean or average effects; statistical assessment of extreme values is more difficult. In military applications, individual differences in design are often handled by presenting design criteria or evaluation results in terms of percentiles. Thus, a workload assessment might be reported in terms of the 5th percentile or 95th percentile values for workload measures obtained. However, extremb behavior or perfo'rmance may reflect more on the individual than on the technology. Thus, caution is urged in looking at extreme behavior when evaluating new technologies.

Tradeoffs Among Device-induced Workload and More General Safety Benefits. Much of what has been discussed has been in the area of device-specific loads. Yet, it is also relevant to consider the broader picture. For example, an In-Vehicle Safety Advisory and Warning System (IVSAWS) might impose a relatively high driver resource loading when used (which might compromise highway safety). On the other hand, IVSAWS provides the driver with critical information about hazards along the route (which might enhance highway safety). At this point, some type of cost-benefit tradeoff would be the most likely means to bring these two discrepant findings together and arrive at a consolidated device assessment.

Multiple Transactions and Tn-cab Device Workload Assessments. Most sophisticated in-cab devices will perform more than one function and so more than one driver transaction is possible. It may well be that, for the same device, one transaction (e.g., menu search) can be accomplished quickly while another transaction (say orienting a moving map display) is cumbersome (but also highly informative once completed). How should these be weighted if one wishes a composite score?

The Chaotic Nature of Crashes. Battelle and its subcontractors recently completed a substantial effort to analyze the major types of crashes that occur in the United States (see Tijerina, 1995, for a synopsis). Analyses were conducted of rear-end crashes, roadway departure crashes, backing crashes, lane-change crashes, various types of intersection crashes, and opposite direction crashes, Based on detailed crash records, the report for each crash type identified putative causal factors and simple kinematic models of crash avoidance requirements. The reports generated from these analyses are intended to support development of crash avoidance systems.

Upon reflection, it appears that while certain causal factors may be attributed to crash incidence as general trends (e.g., driver inattention being a chief causal factor, and hence the motivation behind workload assessment), crash occurrence is in essence a chaotic process.

The word chaotic is used because the. presence of chaos suggests that even if all variables in a non-linear system (the driver/vehicle/driving condition system) could be accounted for, general patterns of system behavior (e.g., crash incidence) may be predicted but specific behaviors (e.g., crash occurrence) may not. In fact, accounting for all variables is presently not possible.

One general finding of the crash problem studies (and other research as well) is that driver inattention is a key contributor to crashes on the highway. Crashes may indeed occur when the driver is not paying attention to the driving scene, but drivers who do not pay attention to their driving do not always have crashes. Crashes occur when a set of circumstances come together in space and time to jointly yield an unfortunate outcome. If drivers are rational within their situational understanding of the driving conditions and their motivations, it is plausible to assume that drivers involved in crashes were inattentive because they expected it to be acceptable to be momentarily inattentive, and their expectations were violated by events at the time. If inattentiveness is risky, then other types of risk-taking (e.g., speeding, following too closely, inappropriate lookout) might also reflect expectancy violations or the mistaken belief that such behaviors will have no adverse outcomes. Given that no one is totally attentive to the driving task at all times, the chaotic nature of crash occurrence illustrates the essence of the phrase "But for the grace of God, there go I."

What does this theory of crashes have to do with safety-relevant workload assessment? Perhaps the best answer is that the possibility of drawing high associations between workload measures and the "ground truth" of highway safety (i.e., crashes) is remote. Instead, the chaotic nature of crash occurrence may be taken to imply that new technology that takes the driver's eyes off the road and attention away from the primary task of driving produces an incremental rise in crash hazard exposure. That is, workload assessment can be used to show that one device increases or decreases crash hazard exposure relative to some other in-vehicle device or transaction. However, it is unlikely in the foreseeable future that precise quantitative predictions of crash incidence will be obtained from workload measures such as those incorporated into the workload assessment protocol. Thus, the state of the art in workload assessment is such as to support relative rather than absolute assessments.

A Workload Safety Paradox. It is known that driver inattention is a leading contributing factor to crashes. As brought out in Task 4, drivers adjust their attention and driving according to their situational understanding of the driving conditions as well as their personal motivations. It is often the case that this situational understanding is faulty, based on a confluence of factors that come together in space and time to lead to a crash. This is why many crashes occur under "ideal" driving conditions, e.g., in daylight and on dry pavement. Thus, it appears that when drivers perceive driving condition demand to be low, they tend to expect nothing out of the ordinary to happen, and elect to reduce their attention to the driving task.

This leads to a workload safety paradox, one that involves at least the three factors of driving condition demand, in-cab device workload demand, and driver discretion to use an in-cab device (see Figure 9.1). If driving condition demand is high (that is, if driving task workload is high), drivers are less likely to devote attention away from the driving task. This was clearly evident in the Task 6 baseline study data. In this case, the workload imposed by an in-cab device becomes relatively less important if the driver can exercise discretion about when or whether to use that device. Of course, a high-workload device still imposes more crash hazard potential than a low-workload device, all else being equal. If, however, device use is mandatory (by operating practice or the nature of the device itself), the relative crash hazard levels increase accordingly because the driver can no longer necessarily choose if or when to interact with the in-cab device.

If driving demand is low (i.e., the driving conditions impose low driving task workload), then the driver is more likely to allocate attention away from the driving task and pursue in-cab device use. This was found in the Task 6 baseline data study and in the Task 7 evaluation of text message and cellular phone use. The relative effects of device workload demand and driver discretion also apply to this driving condition. A high-workload device still imposes more crash hazard potential than a low-workload device, all else being equal. And again, if device use is mandatory rather than discretionary, relative crash hazard levels are elevated. Figure 9.1 depicts these states and indicates the author's hypothetical relative crash hazard level using a rank-order scale from 1 (indicating very low relative crash hazard) to 8 (indicating very high relative crash hazard).

Why would in-cab device use ever be outside the driver's discretion? One possibility may be the operating practices of the fleet for which the heavy vehicle driver works. If, for example, operating practices dictate that a driver is supposed to enter data into a trip recorder within a short time after crossing a state line, this makes such interaction mandatory. A second possibility may be the nature of the device itself. For example, a route guidance or land navigation system may invite the driver to attend to it at junctions (e.g., intersections, merge lanes, and so forth), yet these are areas where traffic conflict potential is greatest. A third possibility is that there may be something intrinsic in human nature that prompts "mandatory" interaction with at least some types of technology. Anyone who has ever compulsively answered the telephone shortly after sitting down to the family dinner, despite the inconvenience, knows this intuitively.

This paradox deserves further research to establish its properties and scope. It suggests, however, that one key aspect of crash avoidance research may involve capture and analysis of driver workload measures to gauge the driver's current attentional state. If crash avoidance system technology (which also overlaps substantially with workload assessment technology needs) can be used to sense and process crash-relevant information, then knowledge of the driver's attentional state could be used to tailor alerts or automatic control intervention accordingly.

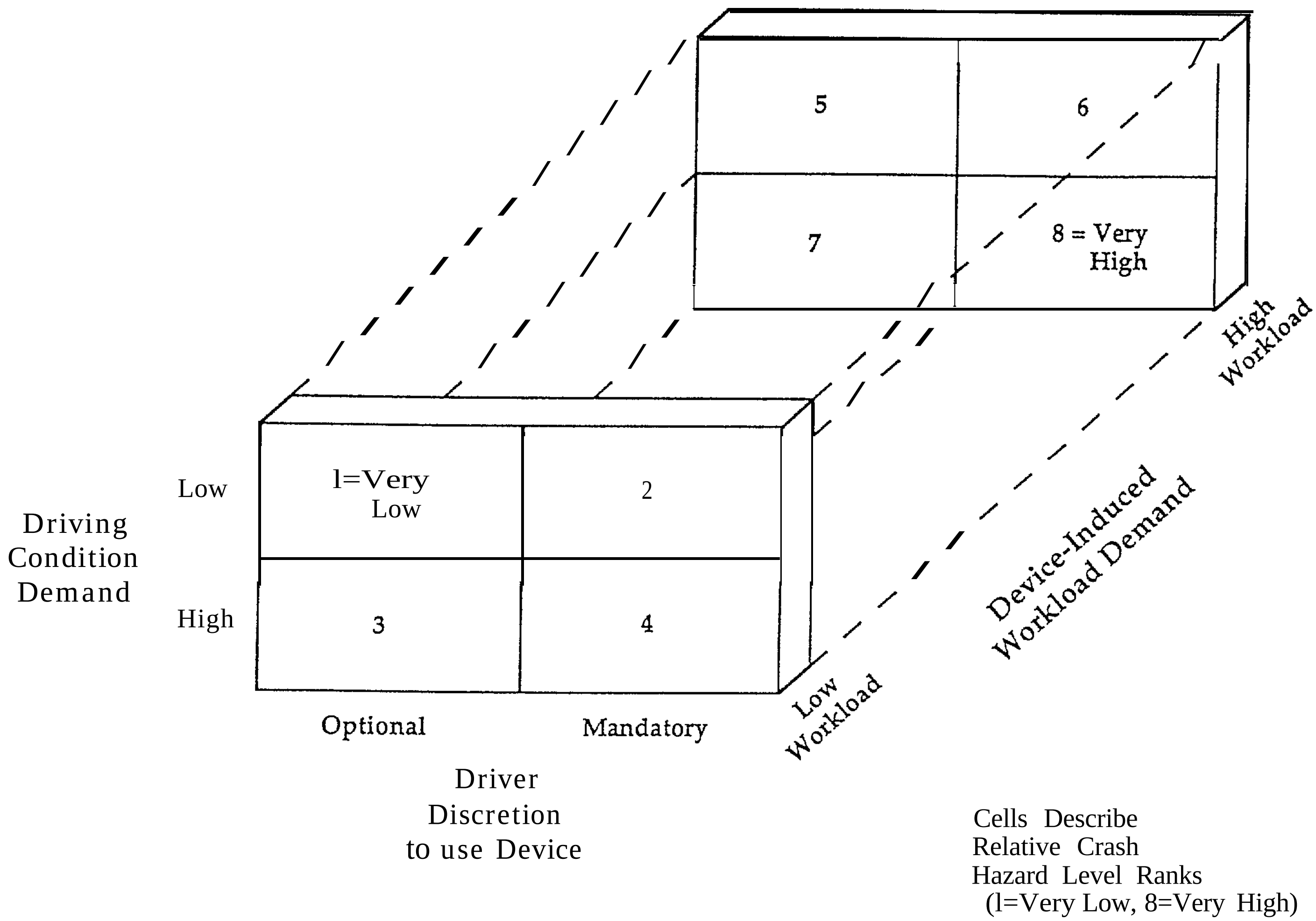

Figure 9.1 Hypothetical Levels of Relative Crash Hazard Given the Interaction of Driving Condition Demand, In-cab Device WorkIoad Demand, and Driver Discretion. (See text for explanation.)

The issue of driver discretion on whether or when to use new technology also deserves further mention. Because of the potential detriment to highway safety, fleets are advised to implement technologies and operational practices that allow the driver maximum flexibility concerning in-cab device use while on the road. This should pay dividends in increased safety and efficiency in the long run.

10.0 RECOMMENDATIONS FOR FUTURE RESEARCH

The workload assessment protocol and research conducted to support its development represents an early step in what is hoped will be a process of continued evolution and improvement of driver-oriented, safety-relevant test and evaluation. This work is most properly considered a beginning rather than an endpoint in terms of device evaluation methodologies. Some fruitful areas for future research to extend the work conducted under this program are provided below.

Further Applications of the Workload Assessment Protocol. Clearly, the utility of the workload assessment protocol and measurement system will depend on the results obtained from applications of it. In particular, the workload assessment methodologies should be applied to Intelligent Transportation System (ITS) products. This will be productive both in terms of refining the workload measurement process and in terms of enhancing the safety and usability of new systems for use in trucks and cars.

Workload Assessment Under Naturalistic Driving Conditions. There is a need to apply the workload protocol under realistic driving. The presence of the on-board experimenter in this program of research likely impacted on the workload measures taken. It should be possible, with emerging data acquisition systems such as DASCAR, to collect data on truly naturalistic driving with minimum intrusion and over an extended period of time.

Collecting Frequencv-of-Use Data on Technologies. Crash hazard potential is a combination of both device-induced workload demand (assessed by the methods and measures in the research protocol) and device frequency-of-use. Frequency-of-use estimates for various in-cab transactions or high technology devices (or analogues) would be useful in furthering an understanding of the safety impacts of existing and new technologies.

Application of the Workload Assessment Protocol to Passenger Vehicles.. The focus in this research program was heavy vehicle drivers and heavy vehicle operation. This focus stemmed from several assumptions. First, it was assumed that commercial operations would likely be the first consumers of new technologies for use in trucks to improve operational efficiency and profitability. Second, heavy vehicle drivers were thought to be more homogeneous with respect to factors such age, driving experience, and training than passenger car drivers at large. This would serve to reduce the potential complications of driver differences in a device evaluation, Third, because heavy vehicle drivers spend considerably more time on the road and travel more miles on the nation's highways than do passenger car drivers, heavy vehicle drivers are exposed to greater crash hazards, all else being equal. Thus, assessment of technologies targeted to heavy vehicle drivers would have potential safety benefits. However, there is a need to extend the application of workload assessment to passenger car drivers. As the price of high technology devices drops, more of these devices will find their way into passenger cars and automobiles driven by a wider range of persons in terms of age, sex, driving experience, and timesharing skills.

The results obtained in this program of research with heavy vehicle drivers are likely to differ from that of passenger car drivers.

Augmenting the Workload Protocol. The heavy vehicle driver workload assessment protocol developed by the Battelle team is intended to be a systematic and thorough guide to the planning and execution of field-oriented test and evaluation of ITS and related high-technology system. It depends on instrumented vehicles, on-the-road data collection, and potentially complex data gathering strategies, especially as one moves to include more research questions into the data collection. A useful adjunct to this approach is one wherein some portion of the assessment could be conducted prior to, or possibly in lieu of, on-the-road empirical assessments. Analysis of device functions, checklist evaluations, and possibly simple video-game simulator assessments might serve as preliminary assessment stages before the more expensive assessments associated with driving simulator and on-the-road field testing were undertaken. This would serve as a screening evaluation, the next step of which is the more rigorous empirical evaluation It appears that there is a need for both types of workload assessment protocol tools. While the current protocol document is well suited to large-scale, well funded ITS demonstrations and field evaluations, it is not currently well suited to the small entrepreneur who may wish to integrate human factors into the design of an ITS product, but cannot muster the resources to execute (or have someone else execute) the current protocol early and perhaps repeatedly in the design process. The earlier in the design process that workload assessments are made, the better the chances that the device design will be revised based on the results of the early-on evaluation.

Development of New Measures and Methods for Workload Assessment. There is a need to expand the array of measures and techniques that were developed under this research program. For example, visual allocation measures were found to be sensitive, robust, and safety-relevant. However, the video camera techniques used in this research program are likely to be too coarse for some device assessments. For example, Head-Up Displays (HUDs) are finding their way into cars and trucks. HUDs may be used for ITS technologies as well as to replace conventional instrument panel displays and controls. While there is great potential for this technology to enhance driver performance and satisfaction, the technology has received only limited assessment for automotive applications. Published studies to date have used simulators and cars, rarely trucks. The safety implications of HUDs represent an important element in bringing this technology to the highway in a coherent fashion. Techniques to analyze HUDs overlap those already developed on this project, but require additional techniques and equipment. In particular, cognitive demand may sometimes lead to a situation wherein the driver “looked but did not see” a crash hazard and failed to react appropriately. Such cognitive “tunnel vision” might be captured by more refined techniques to measure eye movements such as vergence eye movements or near-field accommodation. Pursuit of this option will yield a more comprehensive assessment protocol, one which extends into what may be the primary new display technology of the next 5-10 years.

REFERENCES

Hanowski, R., Kantowitz, B., & Tijerina, L. (September 1995). ***NHTSA heavy vehicle driver workload assessment final report supplement-- Workload assessment of in-cab text message system and cellular phone use by heavy vehicle drivers in a part-task driving simulator*** (Contract No. DTNH22-91-C-07003. Columbus, OH: Battelle.

Kiger, S. M., Rockwell, T. H., Niswonger, S., Tijerina, L,, Myers, L. B., & Nygren, T. (September, 1992). ***NHTSA heavy vehicle driver workload assessment final report supplement --Task 3: Task analysis data collection*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.

Perel, M. (1976). ***Analyzing the role of driver/vehicle incompatibilities in accident causation using police reports*** (Report No. DOT-HS-80 1-858). Washington, DC: National Highway Traffic Safety Administration.

Tijerina, L. (1995). ***Key human factors research needs in IVHS crash avoidance*** (Paper No. 950250). Paper presented at the Transportation Research Board 74th Annual Meeting, January 22-28, 1995, Washington, DC. (Also ***Transportation Research Record,*** in press).

Tijerina, L., Kiger, S. M., Rockwell, T. H., Tornow, C. E., Kinateder, J. G., & Kokkotos, F. (1995, September). ***NHTSA heavy vehicle driver workload assessment final report supplement -- Task 6: Baseline data study.*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.

Tijerina, L, Kiger, S. M., Rockwell, T. H., & Tornow, C. E. (1995, September). ***NHTSA heavy vehicle driver workload assessment final report: Workload assessment of in-cab text message system and cellular phone use by heavy vehicle drivers on the road*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.

Tijerina, L., Kiger, S. M., Rockwell, T. H., & Wierwille, W. W. (1995, June). ***NHTSA heavy vehicle driver workload assessment final report supplement -- Task 5: Heavy vehicle driver workload assessment protocol*** (Contract No. DTNH22-9 l-C-07003). Columbus, OH: Battelle.

Tijerina, L., Kiger, S. M., Wierwille, W. W., Rockwell, T. H., Kantowitz, B., Bittner, Jr., A., Nygren, T., Myers, L., Tolbert, C., & McCallum, M. (January, 1992). ***NHTSA heavy vehicle driver workload assessment final report supplement -- Task 1: Task analysis data and protocols review*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.

Turanski, A., & Tijerina, L. (January, 1992). ***NHTSA heavy vehicle driver workload assessment final report supplement -- Task 2: Standard vehicle configuration/task specifications*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.

Wierwille, W., W., & Tijerina, L. (in press). An analysis of driving accident narratives as a means of determining problems caused by in-vehicle visual allocation and visual workload. To appear in ***Vision in Vehicles V.***

Wierwille, W., W., Tijerina, L., Kiger, S. M., Rockwell, T. H., Lauber, E., & Bittnei, Jr., A. (September, 1992). ***NHTSA heavy vehicle driver workload assessment final report supplement -- Task 4: Review of workload and related research*** (Contract No. DTNH22-91-C-07003). Columbus, OH: Battelle.